A Grammar of Lezgian

Allen Mao

1 Introduction

This paper presents an X' analysis of the Lezgian language, first with X' phrase structure rules that describe a generative grammar for the major lexical categories, e.g. VP (verb phrase), DP (determiner phrase), NP (noun phrase), etc. These rules then build the foundation for a further discussion on question formation, case and agreement, and binding in the Lezgian language in Section 3.

The Lezgian language is spoken in a region of about 5000 km^2 between the Eastern Caucasus mountains and the Caspian Sea, present-day southern Daghestan (Russia) and northern Azerbaijan (16). Haspelmath calculates that with a 90% retention rate and over 466 000 Lezgians in the 1989 census of the Soviet Union, there are well over 400 000 speakers of Lezgian (16). Lezgian is part of the Lezgic branch of the Nakho-Daghestanian language family, commonly known as the "North-East Caucasion" or "East Caucasian" language family (1). Other languages from the Nakho-Daghestanian language family include Chechen and Ingush from the Nakh branch and Avar, an Avaric language (1). Although the economic and political dominance of the Russian language has clearly contributed to the decline of the Lezgian language, Lezgian remains taught at several levels of education in the Republic of Daghestan and Lezgian-language publications, radio broadcasting, and theater productions remain available (24). In addition, Lezgian remains vibrant, if not dominant, in rural areas of the Republic of Daghestan (24). As such, Haspelmath concludes that "as long as the Lezgians remain in their traditional settlement areas, Lezgian is not an endangered language" (24).

As a head-final language, i.e. heads mostly follow complements, the dominant word order of Lezgian is SOV (subject-object-verb), although other word orders do surface especially when spoken (5). Head-finality is compulsory, however, in noun phrases (NP), adjective phrases (AdjP), and postpositional phrases (PP) (5).

Lezgian clauses are uniformly ergative and Lezgian morphology is agglutinative (4-5). There exist 36 cases depending on number (Singular, Plural), case (Absolutive, Ergative, Genetive, Dative, Essive, Elative, Directive), and localization (Ad, Sub, Post, Super, In) (4). Case and localizations occur together, e.g. "Ad" and "Essive" together form the "Adessive" case. Lezgian also exhibits no agreement in noun phrases or finite verbs (6).

Lezgian Examples with Glossing

(1) Absolutive and Gentitive Forms of 1SG pronoun: c.f. Figure 1 for tree

a. Zun ata-na.
I:ABS come-AOR
"I came. (251)

b. [Zi [ğweč'i wax]] ata-na.
[I:GEN [little sister]] came-AOR
"My little sister came. (251)

(2) Selected Verb Forms of *fin* "go" (127). Verb forms do not exhibit agreement. Trees for a, b, c represented in Figure 2

a. fi-zwa
go-IMPF

b. fi-zwa-č
go-IMPF-NEG

c. fi-zwa-j
go-IMPF-PST

d. fe-nwa
go-PRF

(3) Question Types: Polar, Binary-Choice, Constituent, Embdedded Polar, Embedded Constituent, Embedded Alternative:

a. Betxoven.a-n muzyka wa-z k'an-da-ni?
Beethoven-GEN music you-DAT like-FUT-Q
"Do you like Beethoven's music?" (417) (c.f. Figure 3)

b. Professor.di ktab k'el-zawa-j-di j-ani ja taχ̂ajt'a
professor(ERG) book read-IMPF-PTCP-SBST COP-Q or or
kχ̂i-zwa-j-di ja-ni?
write-IMPF-PTCP-SBST COP-Q
"Is the professor reading or writing a book?" (418) (c.f. Figure 4)

c. Farid ata-na-ni?
Farid come-AOR-Q
"Has Farid come?" (7)

d. Farid mus ata-na
Farid when come-AOR
"When did Farid come?" (8)

e. Wun hiχ̂tin dust.uni-n muq'uw mus fe-na?
you:ABS which friend-GEN to when go-AOR
"When did you go to which friend?" (425)

f. Za sadra, kkal.i xa-nwa-t'a, akwa-n.
I:ERG PT [cow(ERG) bear-PRF-COND] see-HORT
"Let me see whether the cow has calved." (425)

g. Jarab abur.u wuč luhu-zwa-t'a?
PT they(ERG) what:ABS say-IMPF-COND
"I wonder what they are saying." (427)

h. Ada-z im axwar ja-ni, xabar ja-ni či-zwa-č-ir.
he-DAT [this:ABS dream COP-Q news COP-Q] know-IMPF-NEG-PST
"He did not know whether this was dream or reality." (426)

(4) Relative clause, Adjective based on preposition: c.f. Figure 6

a. [Kolxoz.di-n jašajiš.di-kaj kx̂e-nwa-j] rasskaz-ar
[kolkhoz-GEN life-SBEL write-PRF-PTCP] story-PL
"stories about collective farm life" (252)
b. revoljucija.di-laj wilikan ümür
revolution-SREL previous life
"life before the revolution" (252)

(5) Demonstratives co-occuring with each other or determiners

a. Ha i klass.d-a sa mus jat'ani za-ni k'el-na-j.
that this class-INESS one when INDF I:ERG-also study-AOR-PST
"At one time I, too, was a student in this classroom." (191)
b. i zi kic'
this I:GEN dog
"this my dog" (not: **dog of this I*) (261)

(6) Postpositions 7

a. Kac stol.di-n k'anik akax̂-na.
cat table-GEN under enter-AOR
"The cat went under the table." (170)
b. Kac stol.di-n k'anikaj xkec'-na.
cat table-GEN from.under go.out-AOR
"The cat came out from under the table." (170)

(7) Ergative and dative subject preceding absolutive argument

a. Alfija.di maq̃ala kx̂e-na
Alfija(ERG) article write
"Alifja wrote an article" (294)
b. Mu?minat.a-z Ibrahim aku-na
Mu'minat-DAT Ibrahim see-AOR
"Mu'minat saw Ibrahim" (294)

(8) Coreferential Omission

a. Ada jarğ-ar.i-z kilig-un dawamar-na.
he(ERG) [distance-PL-DAT look-MSD] continue-AOR
"He kept looking into the distance." (295)
b. Güldeste wiri žüre.di-n k'walax-ar awu-n.i-z mažbur x̂a-na.
Güldeste [all kind-GEN work-PL do-MSD-DAT] forced become-AOR

"Güldeste was forced to do work of all kinds." (295)

c. Wa-z küče.di-z fi-n.i-kaj kiče-zwa-ni?
you-DAT [street-DAT go-MSD-SBEL] afraid-IMPF-Q
"Are you afraid to go on the street?"

Trees

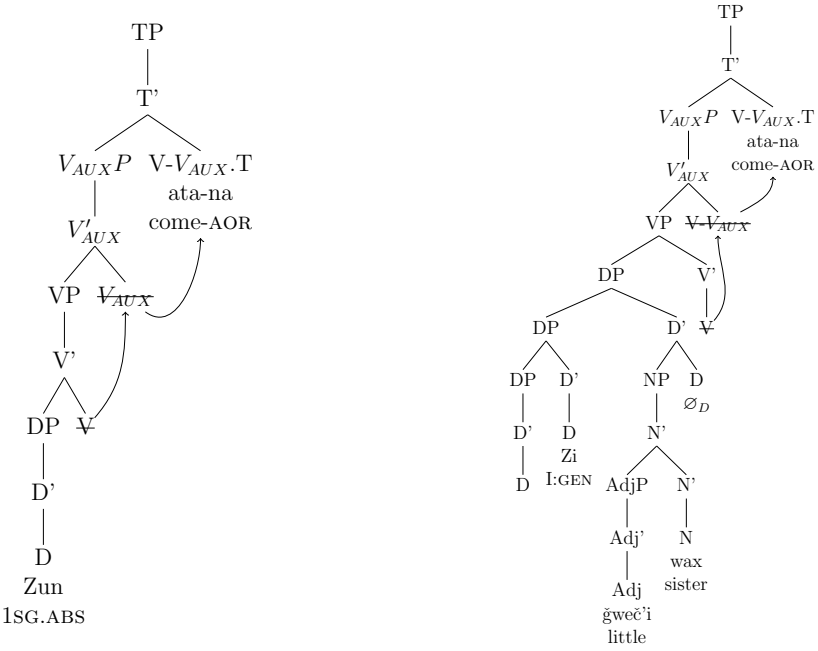

Figure 1: Tree for gloss (1)

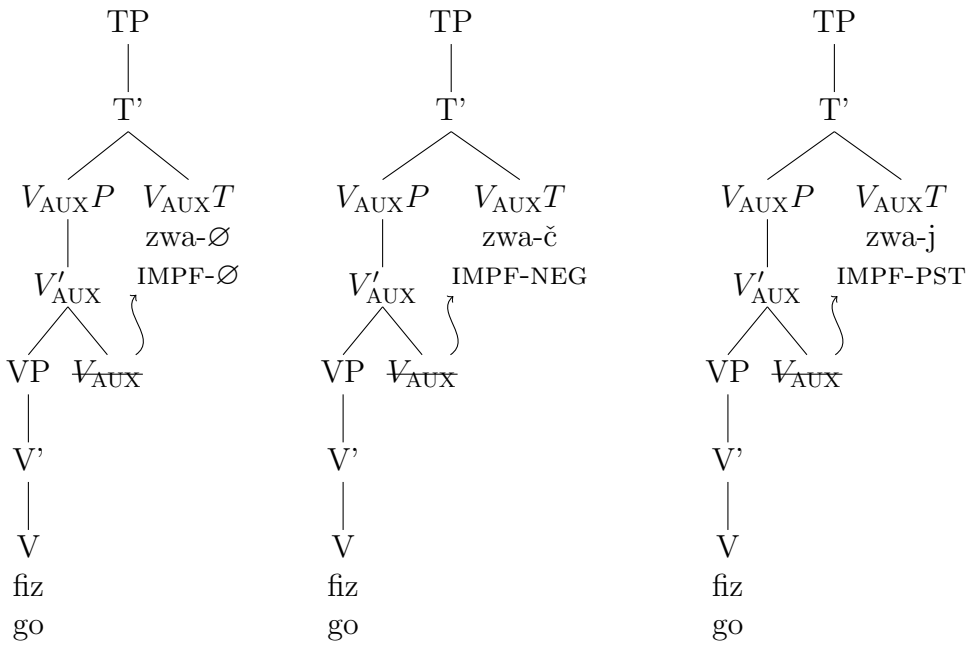

Figure 2: Imperfective affirmative, Imperfective Negative, Past Imperfective Affirmative from gloss (2)

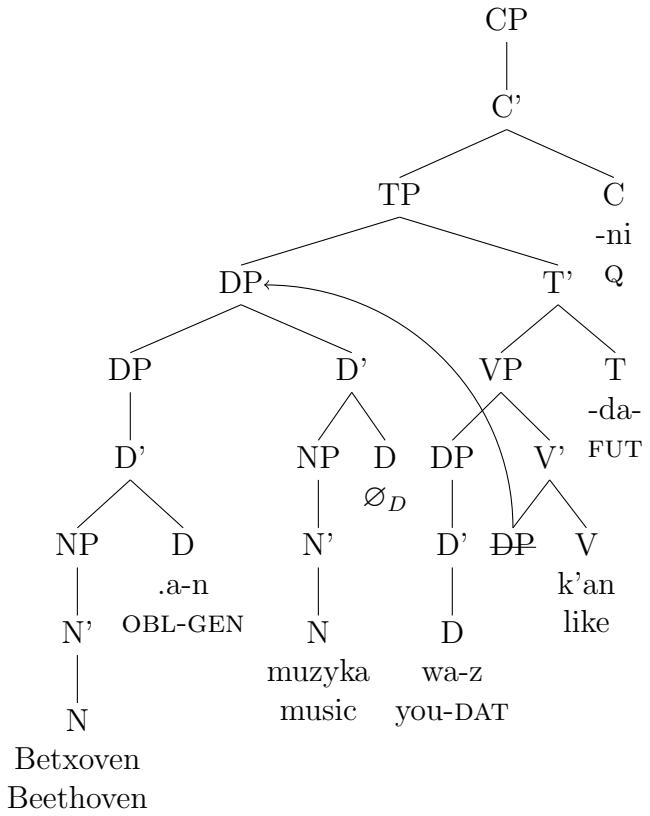

Figure 3: Complete tree for gloss (3-a)

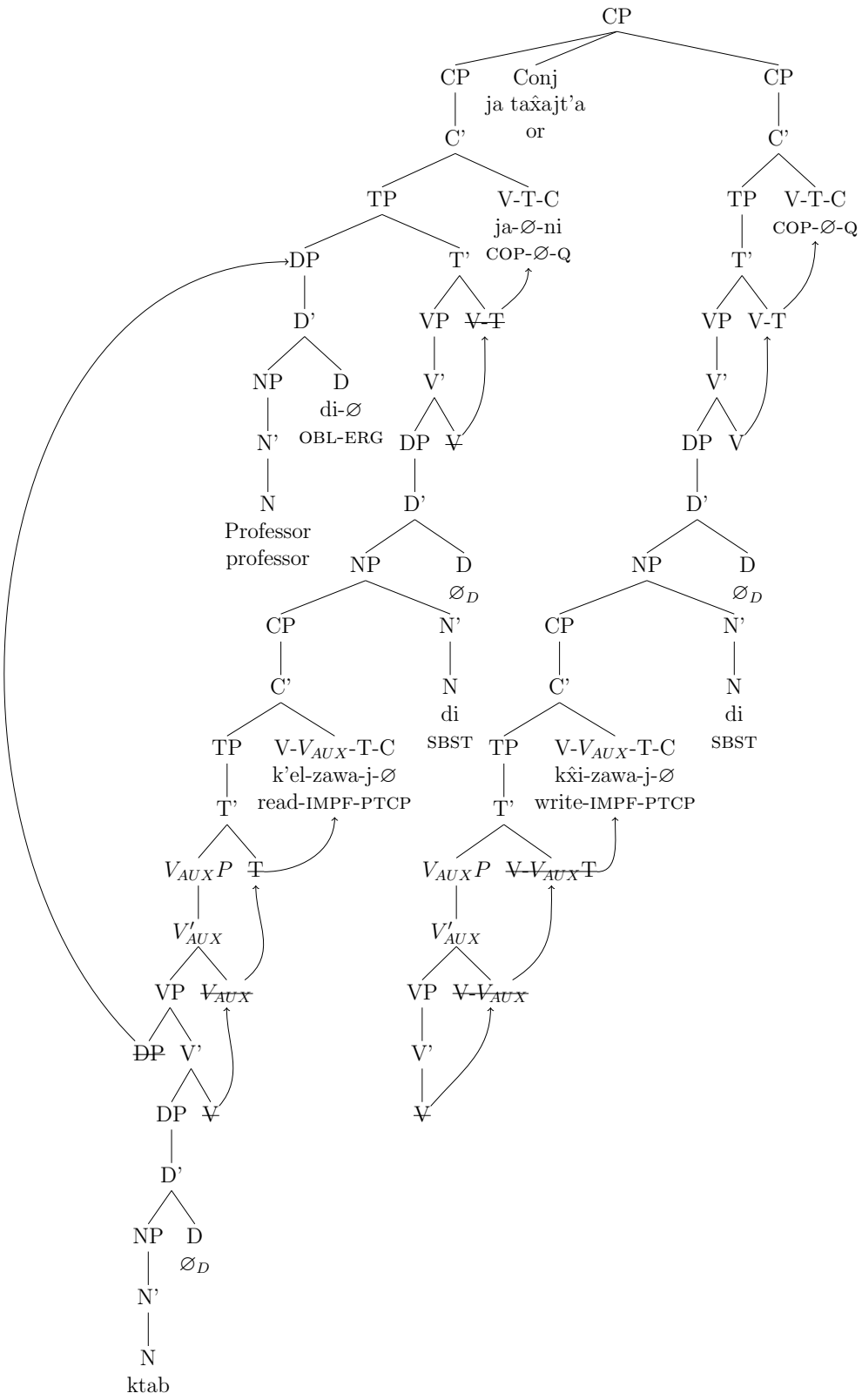

Figure 4: Tree for (3-b)

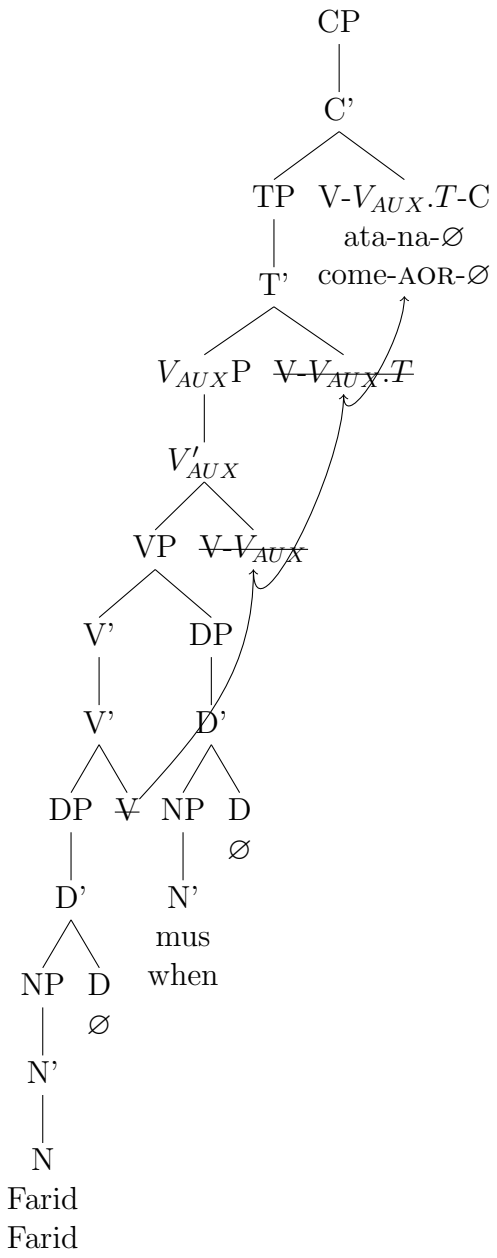

Figure 5: Tree for (3-d)

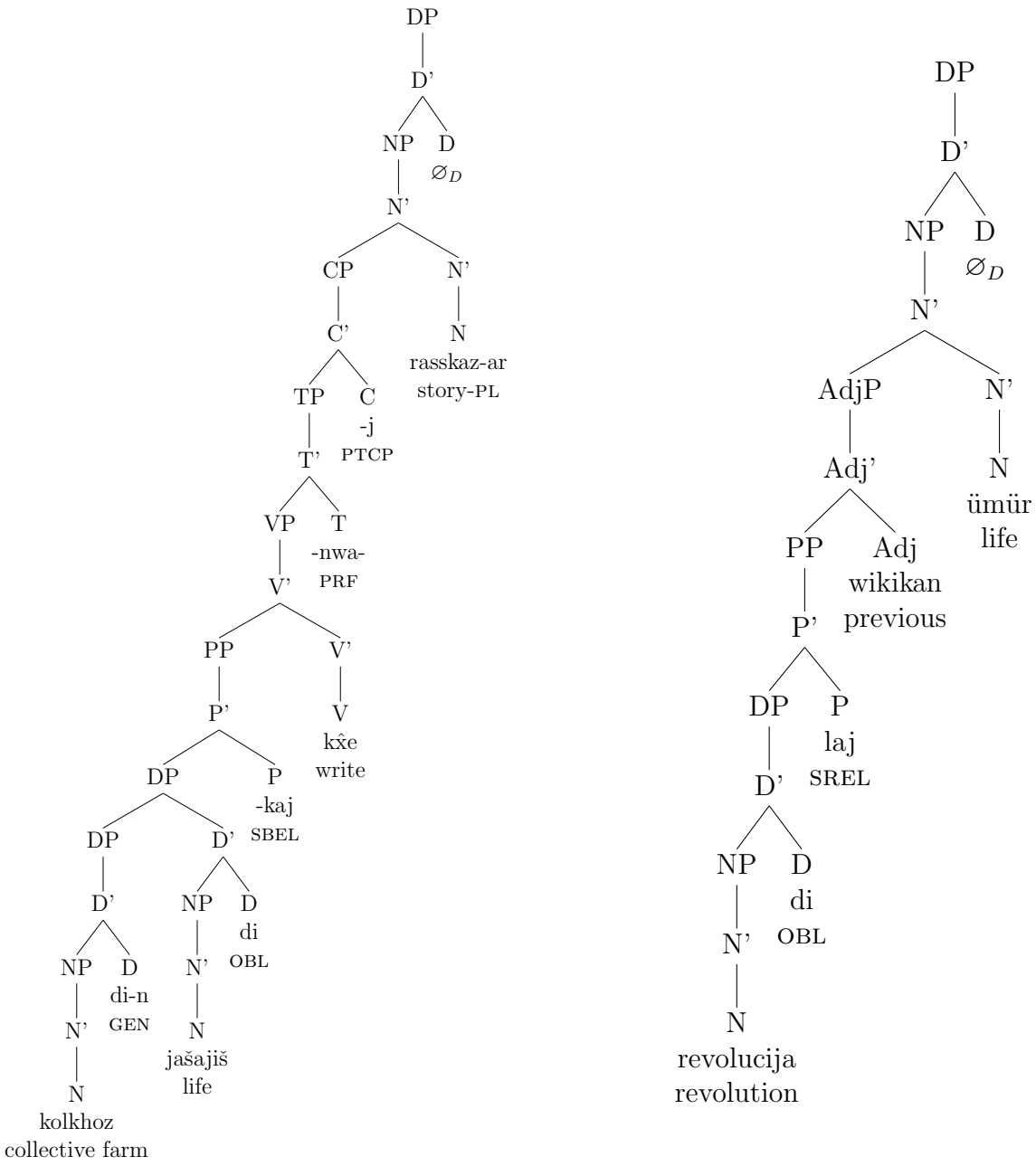

Figure 6: Complete tree for gloss (4-a) and (4-b)

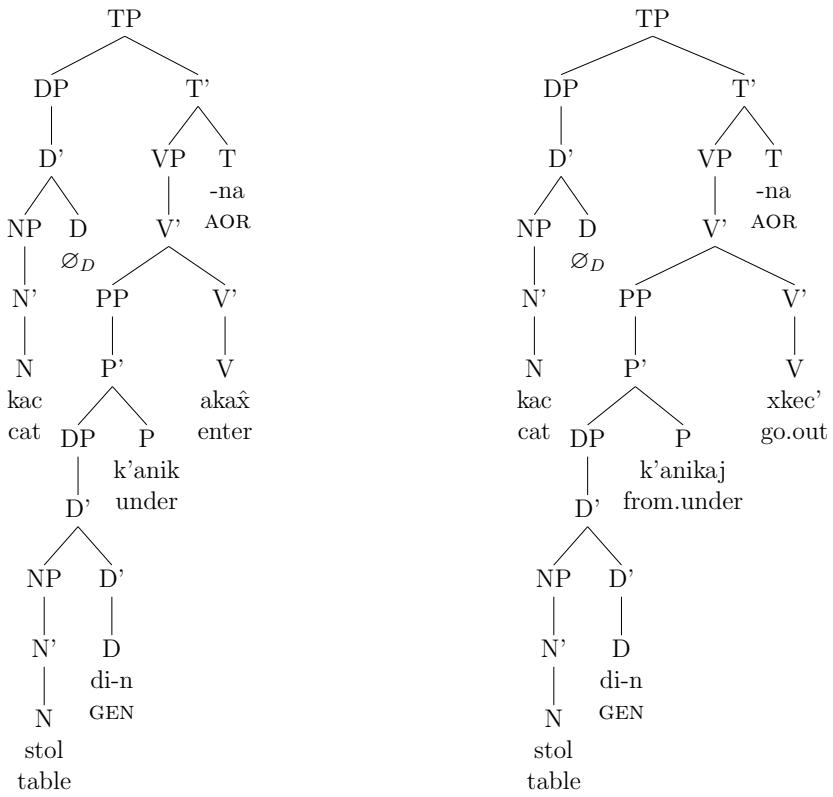

Figure 7: Complete Tree for gloss (6-a) and (6-b)

2 Phrase Structure

2.1 Verb Phrase

All Lezgian verbs take between one and three arguments, inclusive (268-269). As per Lezgian's Ergative-Absolutive paradigm, most Lezgian verb valence patterns contain an absolutive argument (269). This absolutive argument always acts as a theme (268, 270). Each of such intransitive valence patterns also has a corresponding transitive valence pattern of the same arguments plus an Ergative argument (269). This Ergative argument always functions as the agent (270). Lezgian verbs can also accept Locative or Dative arguments, the former serving as recipients and experiencers and the latter as a local case. While there is weak evidence that both Ergative agents and Dative experiencers precede the Absolutive argument, word order is flexible in Lezgian (294-295). Lezgian verbs also lack subject-verb agreement (294), but the free word order suggests head movement.

VP $\rightarrow$V'	(no specifier)
V' $\rightarrow$(DP+)(PP+)(AdvP+)V'	(adjuncts)
V' $\rightarrow$(DP+)(PP+)V	(complements)

2.2 Auxiliary Verb Phrase and Tense Phrase

Lezgian's agglutinative nature makes possible to isolate the affixes that indicate tense and aspect. Refer to the examples in glosses

The phrase structure rules are written as follows:

$$\begin{aligned} \text{TP} &\rightarrow \text{DP T'} \\ \text{T'} &\rightarrow V_{AUX}P \text{ T} \\ V_{AUX}P &\rightarrow V'_{AUX} \\ V'_{AUX} &\rightarrow \text{VP } V_{AUX} \end{aligned}$$

As such, it is possible to distinguish V_{AUX} heads from T heads.

2.3 Subjects

As a role-dominated ergative-absolutive language, the concept of the subject in Lezgian is less straight-forward than it is in a reference-dominated nominative-accusative language like English (294). While ergative agents and dative experiencers may both be subjects as they precede absolutive arguments in unmarked order, the flexible word order of Lezgian makes this evidence very weak (294-295). Instead, the evidence for subjects in Lezgian comes from omissions in coreferential constructions just as they are in English (295). For example, in the English sentence *Maria promised Kim to meet Hans*, the subject of *to meet* and *to promise* corefer to *Maria* (295). In this case, the former serves as the target and the latter as the controller of the omission (295). Just like in English, Lezgian subjects often serve as controllers (295). Example sentence (8-a) shows how the ergative argument behaves as the subject as its omitted in the dependent clause just as example sentences (8-b) and (8-c) do for absolutive and dative arguments.

2.4 Complementizer Phrase

Whereas Lezgian declarative sentences do not exhibit an overt complementizer, polar, i.e. yes-or-no, questions are expressed in Lezgian with the interrogative mood verb affix "-ni" (417). This suffix can be classified as the complementizer in question formation (c.f. gloss (3-a)). Similarly, when a question asks for a selection among a number of options, e.g. *Professor.di ktab k'el-zawa-j-di j-ani ja taxajt'a kxi-zwa-j-di ja-ni?* ("Is the professor reading or writing a book?"), the suffix "-ni" follows both verbs (418); this situation can be modelled as a conjunction phrase joining the two complementizer phrases where each conjunction phrase is headed by the "-ni" suffix (c.f. gloss (3-b)).

Unlike questions in English, Lezgian parametric questions do not move an interrogative phrase to a clause-initial position (421). Instead, the questioned constituent is subsituted with an interrogative pronoun (421). These parametric questions bear the conditional verb suffix *-t'a*. This suffix serves as the complementizer for parametric questions.

Given this information and Lezgian's right-headedness, we write the CP rules as follows:

$$\begin{aligned} \text{CP} &\rightarrow \text{C'} \\ \text{C'} &\rightarrow \text{TP C} \end{aligned}$$

2.5 Noun and Determiner Phrases

Haspelmath describes a noun phrase as one of the following (252):

1. a pronoun
2. a noun head with optional preceding modifiers, i.e. quantifiers, demonstratives, adjective phrases, Genitive NPs, relative clauses
3. a nominalized clause

This paper modifies Haspelmath's categorization to fit the X' schema with the notion of the Determiner Phrase (DP). Pronouns are recategorized as Determiner heads, but unlike the X'-schema for English, quantifiers and demonstratives are categorized as N's.

As modifiers precede the head noun, the X' rules for the NP are as follows:

$$\begin{aligned} \text{NP} &\rightarrow \text{N'} \\ \text{N'} &\rightarrow (\text{AdjP+})\ \text{N'} \\ \text{N'} &\rightarrow \text{N} \end{aligned}$$

Noun phrases in Lezgian do not allow postpositional phrases, NPs in oblique cases, or adverbs as adjunct modifiers (252). As such, English NPs that would normally require a PP adjunct such as "stories about collective farm life" (c.f. gloss (4-a) and (4-b)) would employ a relative clause for the same meaning (252).

Just as in English syntax, pronouns are determiners (c.f. gloss (1). However, as we see in gloss (5-a), since the demonstratives *ha* ("that") and *i* ("this") can form the combination "ha i" (this) (190-191), demonstratives can co-occur and therefore cannot be determiners. This analysis is further supported by gloss (5-b), a construction that would be illegal under English syntax. This gloss reveals that demonstratives are actually of category N', similar to the English "one". Refer to for the tree representations of these glosses. do trees

Lezgian marks indefiniteness but not definiteness, as seen in gloss (5-a). However, what Haspelmath identifies as an optional indefinite article (230), *sa* ("one"), is instead classified as an adjective in this paper. In its place, the oblique case endings fill the role as a determiner head when present. This classification allows for an analysis that preserves Lezgian's right-headedness while assigning a syntactic role for the oblique case markings.

The X' rules for the DP are thus summarized as follows:

$$\begin{aligned} \text{DP} &\rightarrow (\text{DP})\ \text{D'} \\ \text{D'} &\rightarrow (\text{NP})\ \text{D} \end{aligned}$$

As shown in gloss (1), genitive DPs serve as specifiers to the DP they modify.

2.6 Postpositional Phrase

Postpositions are the only type of adposition in Lezgian and are often derived from spatial adverbs, spatial nouns, or converbial verb forms (213). In addition to these postpositions that Haspelmath identifies, this paper will also consider the endings of the locative cases, i.e. essive, elative, and directive with ad, sub, post, super, and in localizations, as prepositions.

An example of the former is shown in the glosses of (6). Glosses (4-a) and (4-b) demonstrate the latter.

The X' rules are summarized as follows:

$$\text{PP} \rightarrow \text{P'}$$
$$\text{P'} \rightarrow \text{DP P}$$

3 Special Topics

3.1 Questions

As mentioned in section 2.4, polar questions in Lezgian are formed with the interrogative verb suffix "-ni", e.g.*Betxovenan muzyka waz k'anda**ni**?* ("Do you like Beethoven's music?") or *Farid atana**ni**?* ("Has Farid come?") (c.f. example (3-a) and (3-c), respectively, for examples glossed) (417). A similar structure is used for questions that ask for a selection among a number of the choices, e.g. *Professordi ktab k'elzawajdi ja**ni** ja taxajt'a kxizwajdi ja**ni**?* ("Is the professor reading or writing a book?") (c.f. example (3-b) for glossing) (418).

Lezgian content (Wh-) questions, on the other hand, do not exhibit any verb affix and instead substitute the constituent in question with the associated interrogative pronoun, e.g. *mus*("when") in *Farid mus atana?* ("When did Farid come?") (c.f. example (3-d) for gloss) (8). One single content question may also contain multiple interrogative pronouns, c.f. example (3-e). A comparison between the tree structures of polar question *Farid atana**ni**?* ("Has Farid come?") in Figure 4 and the content question *Farid mus atana?* ("When did Farid come?") in Figure 5 shows how verb hopping may produce both structures and how a C-head integrates in both polar and content questions.

Embedded questions in Lezgian, both polar, e.g. *Za sadra, kkali xanwa**t'a** akwan*("Let me see whether the cow has calved."), and parametric, e.g. *Jarab aburu wuč luhuzwa**t'a**?*("I wonder what they are saying."), are marked by the conditional suffix *-t'a* on the verb (426). This serves as the C-head for polar and embedded questions as seen in glossed examples (3-f) and (3-g). Embedded choice questions, e.g. *Adaz im axwar ja**ni**, xabar ja**ni** čizwačir*("He did not know whether this was dream or reality") exhibit a similar structure where the C-head "-ni" affixes to both verbs, as glossed in example (3-h) (426).

As seen in these examples, word order in Lezgian is very flexible save its verb finality. While the absolutive argument precedes the dative argument in example (3-a) (tree), in example (3-b) (tree XX), the ergative TP specifier precedes the absolutive argument; few word order restrictions are enforced in Lezgian and the order of arguments and adjuncts in Lezgian clauses is very free (298). This paper accounts for this flexibility with movement, but this very flexibility suggests that this movement is not systematic.

reference tree

3.2 Case/Agreement

As was mentioned in section 1, Lezgian has eighteen cases in total, four grammatical and fourteen local (74). The four grammatical cases are the absolutive, ergative, genitive, and the dative (74). The local cases are divided into five localizations: Ad, Post, Sub, Super, In, each of which has three locatives: Essive, Elative, and Directive (74). As Lezgian is

missing the indirective case, there are only 14 combinations(74). The tables below depict the singular inflections of *sew* ("bear"), first with the grammatical cases, then with the local cases:

Case Name	Lezgian Noun Form	English Translation
Absolutive	sew	the bear
Ergative	sew-re	the bear
Genitive	sew-re-n	of the bear
Dative	sew-re-z	to the bear
Adessive	sew-re-w	at the bear
Adelative	sew-re-w-aj	from the bear
Addirective	sew-re-w-di	toward the bear
Postessive	sew-re-q^h	behind the bear
Postelative	sew-re-q^h-aj	from behind the bear
Postdirective	sew-re-q^h-di	to behind the bear
Subessive	sew-re-k	under the bear
Subelative	sew-re-k-aj	from under the bear
Subdirective	sew-re-k-adi	to under the bear
Superessive	sew-re-l	on the bear
Superelative	sew-re-l-aj	off the bear
Superdirective	sew-re-l-di	onto the bear
Inessive	sew-re	in the bear
Inelative	sew-räj	out of the bear

It is also worth noting the existence of the oblique stem on which all cases but the absolutive are based on (4). For example, in the table of examples above, the oblique stem appears identical to the ergative case, i.e. as *sew-re*; the ergative case is a null suffix attached to the oblique stem (14).

As mentioned in section 1 and seen in the table above, Lezgian is an ergative-absolutive language. However, the flexible word order and lack of agreement in Lezgian makes it difficult to assign morphological rules that determine grammatical case, i.e. Absolutive, Ergative, or Dative, to any particular structural position. For example, as alluded in 2.3, both ergative agents and dative experiencers tend to precede absolutive arguments, thus implying that one of them is likely to jump to the TP specifier position, the flexibility of Lezgian word order also produces examples where the absolutive argument has to jump to the TP specifier position in order to satisfy the X' rules, c.f. Figure 3.

Lezgian's agglutinative morphology, however, simplifies the process of isolating morphemes and, in the context of the locative cases, the analysis of the locative cases as prepositional phrases: except for the inessive and inelative cases, each locative case is marked by a suffix that is added to the oblique form(meaningless form identical to ergative from which all non-absolutie forms are derived). This locative suffix serves as the P head for a prepositional phrase while the oblique suffix acts as a D head in a determiner phrase. This is illustrated in example (4-a) where *-kaj*, the subelative case marker, serves as the P head for the Prepositional Phrase and *di*, the oblique stem, as the D head. Similarly, in example (4-b), *laj*, the superelative case marker, serves as the P head for the prepositional phrase and *di* serves as the D head again. Both of these examples are illustrated in Figure 6.

4 Abbreviations

ABS = absolutive, AOR = aorist, COND = conditional, COP = copula, DAT = dative, ERG = ergative, FUT = future, GEN = genitive, HORT = hortative, IMPF = imperfective, INDF = indefinite, INESS = inessive case, MSD = masdar, NEG = negative, OBL = oblique, PL = plural, PRF = perfect, PST = past, PT = particle, PTCP = participle, Q = question particle, SBEL = subelative case, SBST = substantivizer, SG = singular, SREL = superelative case.

References

[1] Haspelmath, M. (2011). *A grammar of Lezgian (Vol. 9).* Walter de Gruyter.